

# AID can restrict L1 retrotransposition suggesting a dual role in innate and adaptive immunity

Donna A. MacDuff, Zachary L. Demorest and Reuben S. Harris\*

Department of Biochemistry, Molecular Biology and Biophysics, Institute for Molecular Virology, Center for Genome Engineering, University of Minnesota, Minneapolis, MN 55455, USA

Received October 22, 2008; Revised January 8, 2009; Accepted January 9, 2009

## ABSTRACT

**Retrotransposons make up over 40% of the mammalian genome. Some copies are still capable of mobilizing and new insertions promote genetic variation. Several members of the APOBEC3 family of DNA cytosine deaminases function to limit the replication of a variety of retroelements, such as the long-terminal repeat (LTR)-containing MusD and Ty1 elements, and that of the non-LTR retrotransposons, L1 and Alu. However, the APOBEC3 genes are limited to mammalian lineages, whereas retrotransposons are far more widespread. This raises the question of what cellular factors control retroelement transposition in species that lack APOBEC3 genes. A strong phylogenetic case can be made that an ancestral *activation-induced deaminase (AID)*-like gene duplicated and diverged to root the APOBEC3 lineage in mammals. Therefore, we tested the hypothesis that present-day AID proteins possess anti-retroelement activity. We found that AID can inhibit the retrotransposition of L1 through a DNA deamination-independent mechanism. This mechanism may manifest in the cytoplasmic compartment co- or posttranslationally. Together with evidence for AID expression in the ovary, our data combined to suggest that AID has innate immune functions in addition to its integral roles in creating antibody diversity.**

## INTRODUCTION

The formation of highly specific antibodies of multiple isotypes requires the activity of activation-induced deaminase (AID), which deaminates cytosine bases to uracils (C to U) in single-stranded DNA (1–9). At expressed antibody loci, these deamination events trigger somatic hypermutation (SHM) at the immunoglobulin variable

regions and class switch recombination (CSR) at the switch regions. AID and antibody diversification are highly conserved in vertebrates from fish to primates (although fish do not undergo CSR) (10–13).

AID is a member of a much larger family of deaminases that includes the APOBEC3 (A3) proteins, which play a critical role in the innate immune response [for recent reviews see (14,15)]. Many of the A3 proteins can inhibit the replication of a variety of retroviruses. For example, human A3G has potent activity against human immunodeficiency virus (HIV)-1 and Murine Leukemia Virus (MLV), predominantly through C to U deamination of the viral plus-strand cDNA during reverse transcription (16–21).

A number of the A3 proteins have also demonstrated activity against two fundamentally different classes of endogenous retroelement: long-terminal repeat (LTR)-containing retrotransposons, such as MusD of mice and Ty1 of yeast, and non-LTR retroelements, such as long interspersed nucleotide element 1 (LINE1, L1) (22–34). LTR-retrotransposons, which are structurally similar to HIV-1 and other retroviruses, predominantly undergo reverse transcription in the cytoplasm of an infected cell. Inhibition of the LTR-retrotransposons also most likely occurs by DNA deamination during reverse transcription, but deamination-independent mechanisms are also possible (16,17,35). Furthermore, several inactive endogenous retroelements bear strand-specific G-to-A mutational signatures characteristic of A3-dependent hypermutation (36–38). In contrast, the non-LTR retrotransposon L1 undergoes target-primed reverse transcription in the nucleus of the host cell (39,40). Inhibition of L1 retrotransposition by human A3B or A3F does not appear to involve mutation of the retroelement DNA or require A3 catalytic activity (22,23,25,27,31,33,34). However, the anti-L1 activity of A3A requires an intact catalytic site glutamate (E72) (25). Thus, at least two mechanisms may be used by A3s to inhibit the replication of L1.

Retroviruses and endogenous retrotransposons are widely distributed from single cell eukaryotes (e.g. yeast)

\*To whom correspondence should be addressed. Tel: +1 612 624 0457; Fax: +1 612 625 2163; Email: rsh@umn.edu

to complex multicellular organisms, such as humans. However, the *A3* genes are only present in placental mammals (41). Phylogenetic studies have indicated that the first *A3* gene(s) arose from an *AID*-like ancestral gene through a series of duplication and diversification events (41–44). The evolutionary history (and the common enzymatic activity) of this gene family led us to hypothesize that the *A3*s acquired their anti-retroviral activities from the *AID*-like gene from which they evolved (45). Here, we have addressed this question by testing the ability of a panel of representative vertebrate *AID* proteins to inhibit the replication of a diverse set of retroelements: L1, MusD and Ty1. Our data showed that *AID* is able to inhibit the replication of L1 and MusD through a DNA deamination-independent mechanism, but it does not have a significant effect against Ty1.

## MATERIALS AND METHODS

### Sequence alignment and phylogenetic studies

The following *AID* sequences were used: human (NM\_020661.1), pig (BP157753.1), mouse (NM\_009645.2), rat (XM\_001060382), chicken (AJ446140.1), zebrafish (NM\_001008403) [the zebrafish sequence cloned and used in the functional studies had one amino acid substitution (R191Q) from this reference sequence], pufferfish (AY621658) and catfish (AY436507). *AID* amino acid sequences were aligned in ClustalX version 1.83.1 (46). The nucleotide sequences were aligned to the amino acid alignment using PAL2NAL (47). Gaps were deleted from both alignments in JalView (48). Recombination breakpoints were ruled out using GARD (49). Phylip seqboot was used to create bootstraps for the nucleotide sequence alignment and then Phylip dnaml was used to generate 100 unique trees from the bootstrapped sequences (50). Phylip consense was used to create a consensus tree and dnaml was used to add the branch lengths.

### Expression constructs

*pYES3/CT constructs.* Human *AID* cDNA was amplified by PCR from plasmid template [pTrc99A-AID; (4)] using primers 5'-NNG GTA CCG CCA CCA TGG ACA GCC TCT TGA TGA ACC-3' and 5'-NNG GAT CCT CAA AGT CCC AAA GTA CGA AAT G-3'. Pig *AID* cDNA was amplified by PCR from a pig EST (BP157753) (51) with primers 5'-NNG GTA CCG CCA CCA TGG ACA GCC TCC TGA TGA AG-3' and 5'-NNG GAT CCT CAA AGT CCC AAC GTA CGA AAC-3'. Mouse *AID* was amplified by PCR from NOD mouse spleen cDNA using primers 5'-NNG GTA CCG CCA CCA TGG ACA GCC TTC TGA TGA AGC-3' and 5'-NNG GAT CCT CAA AAT CCC AAC ATA CGA AAT G-3'. Rat *AID* was amplified by PCR from rat spleen cDNA using primers 5'-NNG GTA CCG CCA CCA TGG ACA GCC TCT TGA TGA AGC-3' and 5'-NNG GAT CCT CAA AGT CCC AAA ATA CGA AAC-3'. Chicken *AID* was amplified by PCR from DT40 B-cell line (52) cDNA with primers 5'-NNG GTA CCG CCA CCA TGG ACA GCC TCT TGA TG-3' and 5'-NNG GAT CCT CAA AGT CCC AGA GTT TTA AAG-3'. Zebrafish *AID*

was amplified by PCR from whole-zebrafish cDNA using primers 5'-NNG GTA CCG CCA CCA TGA TCT GCA AGC TGG ACA GTG-3' and 5'-NNG GAT CCT CAT AAC CCA AGA AGA GCA AAA AC-3'. Pufferfish *AID* was amplified by PCR from pEscHisf.AID (a gift from Dr M. Nussenzweig, Rockefeller University) with primers 5'-NNG GTA CCG CCA CCA TGA TCA CCA AGC TAG ACA G-3' and 5'-NNG GAT CCT CAG AAT CCG AGG AGC TTA AG-3'. Catfish *AID* was amplified by PCR from pTRE2pur-catfishAID (a gift from Dr B. Magor, University of Alberta) using primers 5'-NNG GTA CCG CCA CCA TGA TGA GCA AGC TGG ACA GT-3' and 5'-NNG GAT CCT TAA AGG CCC AGC AGA GCG AA-3'. All *AID* PCR products were digested with KpnI and BamHI and ligated into pYES3/CT (Invitrogen) and confirmed by sequencing.

*pEGFP-N3 and pcDNA3.1-HA constructs.* *AID* cDNAs were PCR amplified using a common 5'-primer 5'-TAA TAC GAC TCA CTA TAG GG-3' (in pYES3/CT) and a 3'-primers specific to each *AID* cDNA: human 5'-NNN GTC GAC AAG TCC CAA AGT ACG AAA TGC-3'; pig 5'-NNN GTC GAC AAG TCC CAA CGT ACG AAA CGC-3'; mouse 5'-NNN GTC GAC AAA TCC CAA CAT ACG AAA TGC-3'; rat 5'-NNN GTC GAC AAG TCC CAA AAT ACG AAA CGC-3'; chicken 5'-NNN GTC GAC AAG TCC CAG AGT TTT AAA GGC-3'; zebrafish 5'-NNN GTC GAC TAA CCC AAG AAG AGC AAA AAC ATC CC-3'; pufferfish 5'-NNN GTC GAC GAA TCC GAG GAG CTT AAG AGC-3'; catfish 5'-NNN GTC GAC AAG GCC CAG CAG AGC GAA GCC-3'. PCR fragments were digested with HindIII and SalI, and ligated into pEGFP-N3 (Invitrogen). The *AID* cDNAs were then excised with KpnI and SalI and ligated into pcDNA3.1-3xHA (22) digested with KpnI and XhoI. *A3A* was PCR amplified from pTrc99A-A3A-3xHA (25) using primers 5'-NNN NGA GCT CGG TAC CAC CAT GGA AGC CAG CCC AGC-3' and 5'-NNN NGT CGA CCA TCC TTC CGT TTC CCT GAT TCT GGA G-3', digested with KpnI and SalI and ligated into pcDNA3.1-3xHA digested with KpnI and XhoI. The *A3B* and *A3G* expression constructs have been described previously (22,53).

*pEAK8 constructs.* Untagged *AID* cDNAs were cloned into pEAK8 (Edge Biosystems) as HindIII/EcoRI fragments from the pYES3/CT constructs. All HA-tagged *AID* and *A3* cDNAs, as well as HA alone, were cloned into pEAK8 from the corresponding pcDNA3.1-HA plasmids using HindIII and XbaI.

*pTrc99A constructs.* *AID* cDNAs were cloned into pTrc99A as KpnI/BamHI fragments from the corresponding pYES3/CT constructs.

*Site-directed mutagenesis.* Single amino acid variants of human *AID* and *A3A* were generated by QuickChange Site Directed Mutagenesis (Stratagene) using the following primers and their complements: E58Q, 5'-CGG CTG CCA CGT GCA ATT GCT CTT CCT CC-3'; W87A,

5'-CAC CTG GTT CAC CTC CGC GAG CCC CTG CTA CGA C-3'; R24E, 5'-GTC CGC TGG GCT AAG GGT GAG CGT GAG ACC TAC CTG TGC-3'; R112E, 5'-GTC TGA GGA TCT TCA CCG CGG AGC TCT ACT TCT GTG AGG ACC G-3'; C87A, 5'-CAC CTC CTG GAG CCC CGC CTA CGA CTG TGC CCG AC-3'; C90A, 5'-GGA GCC CCT GCT ACG ACG CTG CCC GAC ATG TGG CCG-3'; L44K, 5'-GCT ACA TCC TTT TCA AAG GAC TTT GGT TAT CTT C-3'; D45K, 5'-GTG CTA CAT CCT TTT CAC TGA AGT TTG GTT ATC TTC GC-3'; R74K, 5'-GGG ACC TAG ACC CTG GCA AGT GCT ACC GCG TCA CC-3'; Y88K, 5'-CCT CCT GGA GCC CCT GCA AGGA CTG TGC CCG ACA TGT GGC CG-3'; F109K, 5'-CCT CAG TCT GAG GAT CAA GAC CGC GCG CCT CTA C-3'; E167K, 5'-GGG AAG GGC TGC ATA AAA ATT CAG TTC GTC TC-3'; T27E, 5'-CTA AGG GTC GGC GTG AGG AAT ACC TGT GCT ACG TAG TG-3'; T27A, 5'-CTA AGG GTC GGC GTG AGG CCT ACC TGT GCT ACG TAG TG-3'; S38E, 5'-GTA GTG AAG AGG CGT GAC GAA GCT ACA TCC TTT TCA CTG GAC T-3'; S38A, 5'-GTA GTG AAG AGG CGT GAC GCT GCT ACA TCC TTT TCA CTG GAC-3'; Y48H, 5'-CTG GAC TTT GGT CAT CTT CGC AAT AAG-3'; A3A E72A, 5'-GGC CGC CAT GCG GCC TGC GCT TCT TG-3'; A3A W98A, 5'-GGT CAC TTG GTT CAT CTC CGC GAG CCC CTG CTT CTC CTG GG-3'. Double mutants R24E/R112E, T27A/S38A and T27A/S38E were created sequentially. The AID $\Delta$ C variant has been described previously (54). All constructs were confirmed by sequencing.

#### ***Escherichia coli* mutation assays**

BW310 *E. coli* cells, which are deficient for uracil DNA glycosylase (*UNG*), were transformed with pTrec99A-based AID expression plasmids and grown overnight at 37°C on media containing ampicillin. Four independent colonies were used to inoculate liquid LB cultures containing ampicillin and glucose. Saturated cultures were diluted 10<sup>6</sup>-fold and used to inoculate eight liquid LB cultures containing ampicillin and 1 mM IPTG to induce AID expression. An aliquot of the saturated induced culture was plated onto medium containing 100 µg/ml rifampicin (Rif), and an appropriate dilution was plated onto medium containing ampicillin for a viable cell count. All plates were incubated overnight at 37°C to allow colony formation. Mutation frequencies were calculated as the number of Rif-resistant colonies per viable cell. Full procedures were described previously (4,55).

#### **Cell culture and microscopy**

HeLa cells and HEK293 cells were maintained in Dulbecco's modified Eagle medium (Invitrogen) supplemented with 10% fetal bovine serum, 100 U/ml penicillin and 100 µg/ml streptomycin. For microscopy, HeLa cells were seeded in eight-well chambered coverglasses (Nunc LabTek) at 8000 cells per well. 24 h later, cells were transfected with 200 ng of pEGFP-N3-based plasmid DNAs. Twenty-four hours post-transfection, half of the wells were treated with 20 ng/ml leptomycin B

(LC Laboratories) for 2 h. The growth medium was then replaced with phenol-red-free and serum-free DMEM and the cells were imaged using a Zeiss Axiovert 200 microscope (with kind permission from Dr M. Titus).

#### **MusD transposition assays**

HeLa cells were plated in six-well plates at  $1.5 \times 10^5$  cells per well. Twenty-four hours later, cells were transfected with 0.5 µg MusD plasmid, 0.4 µg of an EGFP expression plasmid (pEAK8-GFP) and 1 µg of the AID-HA, APOBEC3G-HA expression plasmid or empty vector (pEAK8-HA) using *TransIT-LT1* (Mirus). For the titration of human AID, AID<sub>E58Q</sub> and A3G against MusD, various amounts (0.5, 1.0 or 1.5 µg) of the AID or APOBEC3G expression plasmid and/or empty vector were transfected into cells (2.4 µg total DNA per transfection). Two days posttransfection, 10% of the cells were analyzed for EGFP expression by flow cytometry and the rest plated into 100-mm dishes in the presence of 850 µg/ml G418. After 14 days, G418-resistant colonies were fixed, stained with crystal violet solution and counted.

#### **L1 retrotransposition assays**

HEK293 cells were plated in six-well plates at  $2 \times 10^5$  cells per well. Twenty-four hours later, cells were transfected with 0.5 µg of the L1 (or control plasmid) and 0.5 µg of the AID-HA, APOBEC3-HA expression plasmid or empty vector (pcDNA3.1-HA) using *TransIT-LT1*-mediated transfection. For the titration of wild-type and mutant deaminases against L1, various amounts (0.25, 0.5 or 1 µg) of the pcDNA3.1-HA-based AID or APOBEC3 expression plasmid and/or empty vector were transfected into cells (1.5 µg total DNA per transfection). Twenty-four hours posttransfection, cells that had received the L1 plasmid were selected for with 0.75 µg/ml puromycin. Three days posttransfection, 35% of the cells were analyzed for GFP expression by flow cytometry, 35% of the cells were lysed for western blot analysis and the remaining 30% of cells were left in the plate in medium containing 0.75 µg/ml puromycin. On day five posttransfection, the remaining cells were analyzed for GFP expression by flow cytometry. Although some inter-experiment variation was observed for transposition frequencies (due to any number of factors), the relative fold-effects attributable to AID/A3 proteins were reproducible and statistically significant (e.g. Supplementary Figure S1).

#### **Ty1 retrotransposition and yeast mutation assays**

Endogenous retrotransposition assays were performed with *Saccharomyces cerevisiae* strain DG1141 (*mat $\alpha$  trp1-hisG ura3-167 his3 $\Delta$ 200 Ty1-2y2his3AI*). DG1141 was transformed with pYES3/CT and its derivatives using a standard lithium acetate heat-shock procedure. Transformants were selected on synthetic complete media lacking tryptophan (SC-TRP). Single colonies were used to inoculate SC-TRP + glucose liquid cultures and incubated at 37°C. Saturated cultures were diluted 500-fold and used to inoculate 2 ml SC-TRP + 3% galactose, which were grown to saturation at 20°C (~10 days). To determine the transposition frequency, 200 µl of each

culture was plated onto SC-HIS media and 100  $\mu$ l of a  $10^5$  dilution was plated onto YPAD (rich) media. The transposition frequency of Ty1 was calculated as the number of HIS<sup>+</sup> colonies divided by the number of viable cells.

Yeast-based canavanine-resistance assays for DNA mutation were performed in a UNG-deficient derivative of the *S. cerevisiae* strain, L40 [*MATa his3 $\Delta$ 200 trp1-901 leu2-3112 ade2 LYS2(4lexAop-HIS3) URA3(8lexAop-lacZ) GAL4*; (26)]. L40-UNG<sup>-</sup> was transformed with pYES3/CT and its derivatives and selected on SC-TRP media. Single colonies were used to inoculate 2 ml SC-TRP + 2% galactose + 1% raffinose liquid cultures, which were grown to saturation at 30°C. Two hundred microliters of each culture was plated onto SC media containing 30  $\mu$ g/ml of canavanine, and 100  $\mu$ l of a  $10^5$  dilution was plated onto YPAD (rich) media. The mutation frequency was calculated as the number of canavanine-resistant colonies per viable cell.

### Immunoblots

HeLa and 293 cells were lysed in 2 $\times$  Laemmli buffer and boiled for 5 min to obtain whole-cell extracts. Proteins were separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), transferred to polyvinylidene difluoride (PVDF) membranes and probed with anti-HA (Covance) or anti- $\alpha$ -tubulin (Covance) antibodies.

### Quantitative PCR analyses

Tissues were harvested from three male and two female C57/BL6 mice and flash frozen in liquid nitrogen. RNA was extracted from the frozen tissue using an RNeasy Mini Kit (Qiagen) and a rotor-stator homogenizer. RNA extracts were treated with DNase (Ambion DNA-free Turbo DNase) and 500 ng of each RNA sample was reverse transcribed using Transcriptor Reverse Transcriptase (Roche) to make cDNA. cDNA samples were further diluted in water and the cDNA equivalent of 10 ng of RNA was used for quantitative PCR (qPCR) in a Lightcycler 480 II (Roche). Primer-probe sets for qPCR for each gene were designed using the Roche Universal Probe Library (UPL) website. *AID* cDNA was detected using primers 5'-TCC TGC TCA CTG GAC TTC G-3' and 5'-GCG TAG GAA CAA CAA TTC CAC-3' and UPL probe 71 (5'-CTGGCTGC); *A1* cDNA was detected using primers 5'-CCC TTG AAA TCA GAT CAG GAA-3' and 5'-CCT GGC TCA TGA TGT CCT C-3' and UPL probe 68 (5'-AGGAGCAG); *A2* cDNA was detected using primers 5'-GGA GAA GTT GGC AGA CAT CC-3' and 5'-TCT GAG TGG CAG CAG GTA AA-3' and UPL probe 97 (5'-TGGAA GTC); *A3* cDNA was detected using primers 5'-TAC CAG CTG GAG CAG TTC AA-3' and 5'-CTG CAT GCT GTT TGC CTT T-3' and UPL probe 27 (5'-GCT GCCTG); and *RPL13A* cDNA was detected using primers 5'-ATC CCT CCA CCC TAT GAC AA-3' and 5'-GCC CCA GGT AAG CAA ACT T-3' and UPL probe 108 (5'-GAGAGCAG). Relative quantitative analysis was performed using Lightcycler 480 SW 1.5 software (Roche).

## RESULTS

### AID proteins from eight vertebrate species are functional DNA mutators in *E. coli*

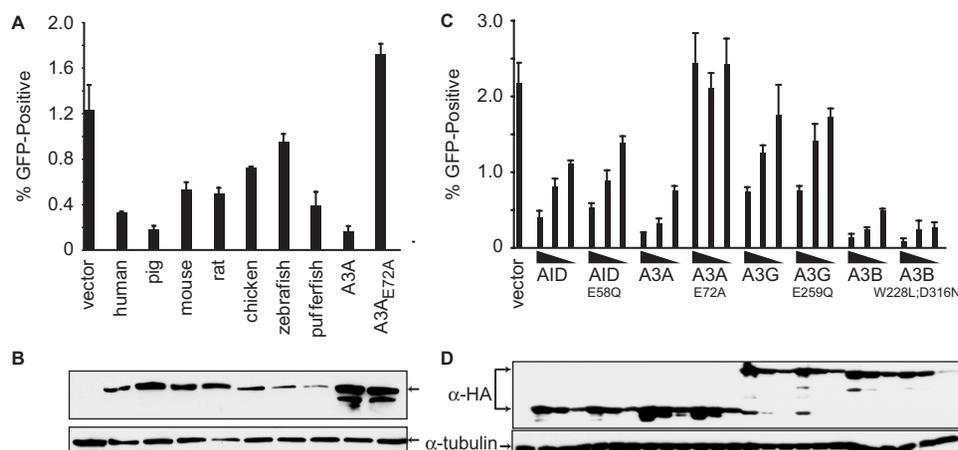
We hypothesized that present day AID (either mammalian or non-mammalian) has the ability to inhibit the replication of retroelements, similar to that of the related APOBEC3 (A3) proteins. Towards testing this hypothesis, we acquired and (sub)cloned *AID* cDNAs from eight representative vertebrate species (Figures 1A and B) and confirmed their ability to mutate DNA using an *E. coli*-based DNA mutation assay (Figure 1C and Supplementary Figure S2). Expression of all of the *AID* proteins in *E. coli* increased the frequency of occurrence of Rifampicin-resistant (Rif<sup>R</sup>) *E. coli* colonies, ranging from 2.2-fold (pufferfish AID) to 12.4-fold (zebrafish AID) over the empty vector. Almost all of these proteins have been shown to elicit activity in *E. coli* previously and, as discussed in these studies, the variation in the mutation frequencies observed here may be due to differences in protein expression, optimal temperatures for enzymatic activity, intrinsic enzymatic activity and/or other factors (43,56–58). Regardless of the multiple explanations, the most important points from these initial studies were that all of these *AID* cDNAs were capable of encoding a catalytically active protein and that they were therefore suitable for subsequent experimentation.

### AID can inhibit the transposition of the non-LTR retrotransposon, L1

Several human A3 proteins have been shown to possess the capacity to inhibit the transposition of L1 in cell culture-based assays (22,23,25,31–34). L1 undergoes target-primed reverse transcription to produce a DNA copy of its genome within the nucleus of the host cell. Since AID is capable of entering the nuclear compartment, we considered that AID might also possess anti-L1 activity. We therefore tested its activity against an L1 reporter construct in 293 cells. The L1 construct used in our studies harbors a *GFP* reporter gene in the reverse orientation to the L1 genome (59). The *GFP* gene contains an anti-sense intron that must be removed by splicing of the full-length L1 transcript, followed by reverse transcription and integration, in order for the GFP protein to be functionally expressed. A GFP-positive cell therefore indicates that a successful retrotransposition event has occurred.

The 293 cells were co-transfected with AID-HA expression plasmids and the L1 reporter plasmid, and transfectants were selected with puromycin (L1 plasmid). L1 retrotransposition was monitored by the appearance of GFP-positive cells 5 days later. All of the *AID* proteins tested inhibited L1 retrotransposition (the appearance of GFP<sup>+</sup> cells) (Figure 2A). The ability of pig AID to inhibit L1 transposition was comparable with the positive control, human A3A [7.3-fold and 8.0-fold, respectively (23,25,31,32)]. Human AID displayed a 3.8-fold inhibition and, pufferfish AID showed a 3.2-fold inhibition despite its lower expression level (Figure 2B) and activity in *E. coli* (Figure 1C and Supplementary Figure S2). Chicken and zebrafish AIDs had the weakest activity against L1 at





**Figure 2.** Inhibition of L1 retrotransposition by AID and A3 proteins. (A) Percentage of the drug-resistant cells that were GFP-positive 5 days after transfection with the L1 and indicated AID/A3A plasmids. Bars represent the mean of three independent transfections. Error bars represent the standard deviation. (B) Western blot showing expression of the HA-tagged AID/A3 proteins from a representative experiment from (A). Tubulin is a loading control. (C) Comparison of the percentage of the drug-resistant cells that were GFP-positive 5 days after transfection with L1 and varying amounts of the indicated AID and A3 plasmids and their corresponding catalytically inactive mutants. The amount of AID/A3 plasmid transfected is indicated below the graph (1.0, 0.5 or 0.25 μg). The histogram bars represent the mean of three independent cultures, and the standard deviation is shown. (D) Western blot analysis of the protein expression levels from the experiment in (C). Tubulin is a loading control.

mechanism is most likely not through mutation of the L1 genome.

To confirm and extend these studies, we tested several mutants of human AID that were expected to disrupt catalysis by preventing zinc coordination (C87A, C90A) or altering the catalytic site pocket [W84A; extrapolated from the NMR structure of the C-terminal domain of A3G (60)]. These mutants also displayed close to wild-type levels of activity against L1, despite being catalytically inactive in the *E. coli*-based DNA mutation assay (Figure 3 and Supplementary Figure S2). In contrast, consistent with the requirement of the integrity of the catalytic site of A3A for L1 restriction, mutation of W98 (equivalent to W84 in AID) to alanine completely eliminated the anti-L1 activity of A3A (Figure 3). These data indicated that AID and A3A restrict L1 replication by two distinct mechanisms.

However, despite these differential mechanistic requirements, L1 inhibition by A3A (or, understandably, AID) does not appear to be associated with hypermutations (22,23,31,32). At present, it is difficult to fully reconcile these observations, but one explanation is the possibility that an edited (uracilated) L1 replication intermediate attracts DNA repair proteins that trigger its rapid degradation.

As an additional control to rule out the possibility that the reduction in GFP expression observed in the presence of AID or A3A may have been due to a nonspecific effect on the L1 plasmid, we repeated the experiment with a selected number of AID and A3 constructs, but replaced the L1 construct with a plasmid that expresses GFP constitutively (pEGFP-N3) (Supplementary Figure S3). Although we detected a decrease of ~10–20% in the percentage of GFP-positive cells in the presence of wild-type AID or A3A, it was clearly insufficient to fully explain the observed reduction in L1 retrotransposition.

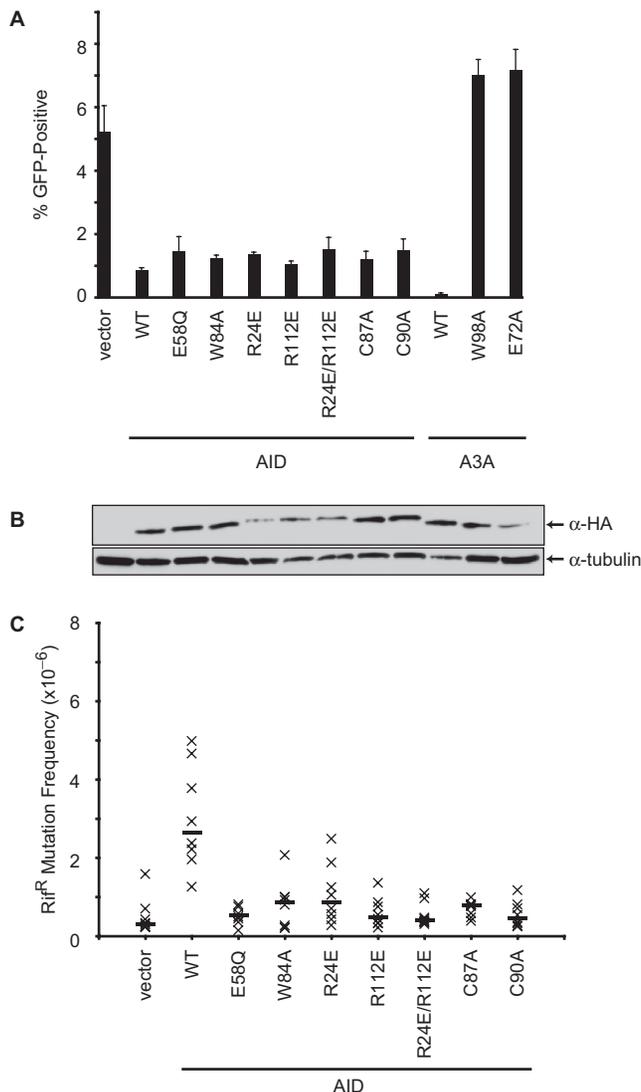
Moreover, this effect appeared dependent upon the catalytic activity of both proteins. A3G did not have any effect on the percentage of GFP-positive cells.

#### DNA binding by AID is not required for the inhibition of L1 transposition

Since the catalytic glutamate (E58) and the active site tryptophan (W84) were not necessary for AID to inhibit L1 transposition, we sought next to dissect this deamination-independent mechanism. First, we asked whether DNA binding was required. Based on the NMR structure of the C-terminal catalytic domain of A3G, we mutated two conserved arginine residues in human AID that are predicted to be required for DNA binding (R24E, R112E and R24E/R112E) (60). The ability of these arginine mutants to impede L1 transposition was almost identical to that of the wild-type protein, suggesting that direct DNA binding is not required for L1 restriction (Figure 3A). Despite some variability in the immunoblot results, all AID variants were expressed similarly (Figure 3B and data not shown). As expected, these AID mutants were inactive in the *E. coli*-based DNA mutation assay (Figure 3C and Supplementary Figure S2).

#### Inhibition of L1 transposition is unaffected by the nuclear accumulation of AID

Since L1 undergoes reverse transcription in the nucleus of the host cell, we initially hypothesized that AID would need to be in the nucleus to inhibit this process. However, since mutation of the L1 DNA by AID is not required to inhibit L1 replication, AID could be acting either within the nucleus or within the cytoplasm, where it is localized predominantly [(61); Figure 4D]. To address this question, we tested a variant of AID that is mostly nuclear because it lacks the last 10 amino acids, which includes the nuclear export sequence (AIDΔC)



**Figure 3.** Effect of AID catalytic and DNA-binding mutants on L1 retrotransposition. (A) Percentage of puromycin-resistant cells that were GFP-positive 5 days after transfection with the L1 and indicated AID variants. A3A and A3A<sub>E72A</sub> were included as controls. Histogram bars represent the mean of three independent cultures, and the standard deviation is shown. WT, wild-type. (B) Western blot showing expression of the HA-tagged proteins from a representative experiment from (A). Tubulin is a loading control. (C) *E. coli*-based Rif<sup>R</sup> mutation assay. Each 'x' represents the mutation frequency of an independent culture, calculated as the number of rifampicin-resistant colonies per viable cell. Eight independent cultures were assayed for each AID variant, and the median mutation frequencies are indicated by the horizontal bars. Vector represents the background level of mutation.

[(62,63); Figure 4D]. AID $\Delta$ C inhibited L1 to the same extent as the wild-type protein (Figure 4).

However, we and others have noted that AID is imported into the nucleus of 293 cells at a lower rate compared with other cell lines (A. M. Patenaude and J. Di Noia, personal communication). We therefore repeated the experiment in HeLa cells. Again, we observed that AID and AID $\Delta$ C inhibited L1 to similar extents despite their clearly distinct subcellular localizations (Figure 4D and Supplementary Figure S4). This result

suggested either that the amount of AID protein required for inhibition of L1 is not limiting in our systems (i.e. the small amounts of protein present in the nucleus or cytoplasm are sufficient to inhibit L1) or that AID inhibits L1 soon after it is translated in the cytoplasm. We also noted that the rate of L1 transposition and the expression of AID are both lower in HeLa cells than in 293 cells (Supplementary Figure S4 and unpublished data). The reason(s) for these differences were not investigated.

#### Phosphorylation of AID does not affect its ability to inhibit L1

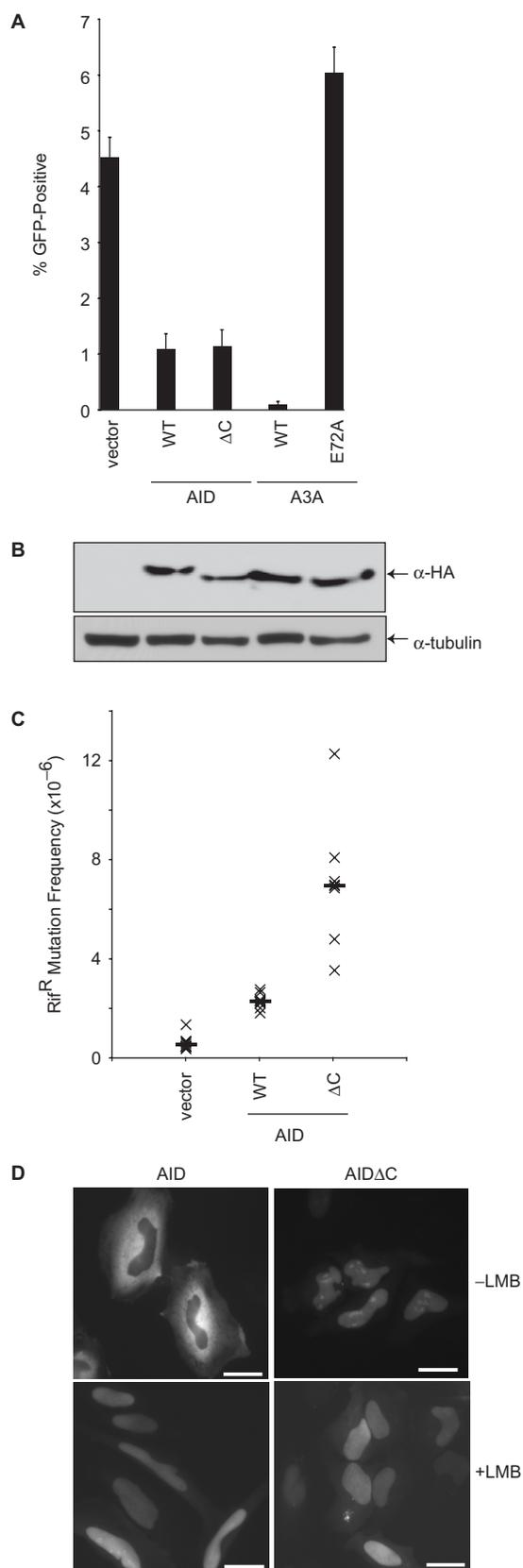
It has been demonstrated previously that AID can be phosphorylated at threonine (T)27 and serine (S)38 and that these events may be important for the function of AID (64–68). Therefore, we considered that phosphorylation could regulate the ability of AID to inhibit L1 transposition, and generated variants of AID with these sites mutated to either alanine (to prevent phosphorylation) or to glutamate (to mimic phosphorylation). Again, these mutants inhibited L1 replication to the same extent as the wild-type protein, indicating that phosphorylation of AID is not required for the restriction of L1 (Figure 5).

#### AID cDNA and RNA are not sufficient to inhibit L1 transposition

Since all of the AID variants we tested behaved almost identically to the wild-type protein, we considered that the anti-L1 activity might be conferred by the AID expression plasmid or the RNA, rather than by the protein. To address this possibility we generated two AID expression constructs that contained early premature termination codons (K10X and W20X). Neither of these constructs was able to inhibit L1 transposition, strongly indicating the requirement for the AID protein (Supplementary Figure S6).

#### AID can also inhibit the transposition of MusD

In addition to inhibiting L1, a non-LTR retrotransposon, A3 proteins are also capable of interfering with the replication of LTR-type endogenous retroelements, such as MusD (16,26–28,32,35–38,69–72). We considered that AID may possess the same activity and therefore tested AID's ability to inhibit MusD transposition. Retrotransposition of a marked MusD element from a plasmid in HeLa cells can be monitored by the appearance of neomycin-resistant (Neo<sup>R</sup>) colonies (73). Consistent with a previously published report, AID proteins from multiple species had a modest, inhibitory effect on the formation of Neo<sup>R</sup> colonies, whereas human A3G almost completely eliminated retrotransposition [(35); Figure 6A]. Analysis of data from four independent experiments confirmed that human AID exerted a significant inhibitory effect against MusD ( $P < 0.0001$ , Student's *t*-test; Supplementary Figure S5). However, we noticed that the expression of AID was ~20-fold lower than that of A3G in HeLa cells, making a direct comparison between the two proteins difficult. To extend this study and to determine whether the inhibitory effect of AID was dependent upon its catalytic activity, we



**Figure 4.** Effect of a nuclear export-defective AID mutant on L1 retrotransposition. (A) Percentage of puromycin-resistant cells that were GFP-positive 5 days after transfection with the L1 and indicated AID variants. A3A and A3A<sub>E72A</sub> were included as controls.

transfected varying amounts of human AID and AID<sub>E58Q</sub> into the cells along with the MusD-containing plasmid and monitored retrotransposition. Again, wild-type AID displayed weak anti-MusD activity that varied with the amount of transfected AID plasmid (Supplementary Figure S3A). Similar to our L1 studies, AID<sub>E58Q</sub> inhibited MusD to the same extent as the wild-type protein (Figure 6C), suggesting a deamination-independent mechanism.

#### AID does not significantly inhibit the transposition of Ty1 in yeast

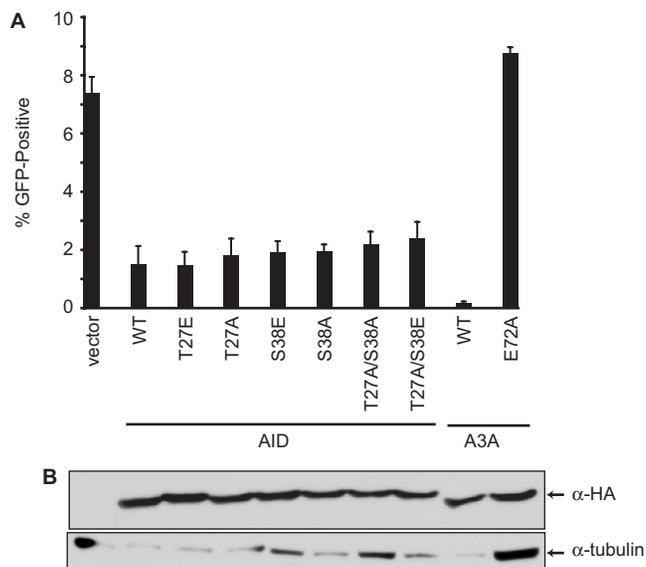
We considered that HeLa and 293 cells (non-B cells) may contain an inhibitor of AID that is interfering with its function. We therefore opted to use an extremely heterologous system: transposition of a yeast LTR-retroelement, Ty1 (26,28). Retrotransposition of the Ty1 genome results in the removal of an intron from a *HIS3* reporter gene, conferring a *HIS*<sup>+</sup> phenotype to yeast cells with a newly retrotransposed Ty1 element. The replication of Ty1, and therefore the formation of *HIS*<sup>+</sup> colonies, was largely unaffected by the presence of any of the AID proteins, in contrast to human A3G and A3F [Figure 7A; (26,28)].

This negative result prompted us to determine whether our AID constructs were capable of catalyzing DNA cytosine to uracil deamination in yeast cells. We determined that all of the AID proteins used increased the rate of mutation to canavanine resistance compared with vector (Figure 7B), indicating that they were all able to mutate the *CAN1* gene DNA, as shown previously for human AID (74). Thus, these controls helped support the conclusion that AID is not able to inhibit the transposition of Ty1 in yeast.

#### AID is expressed outside of the B-cell compartment in mice

While the anti-L1 and anti-MusD activities of AID were certainly intriguing, we were unsure how a protein that is expressed in B lymphocytes could prevent the transposition of retroelements that replicate mostly in germ cells and/or early in embryogenesis (although reports of somatic mobilization of L1 are emerging) (75–78). However, several studies have suggested that AID may be expressed outside of the B-cell compartment in certain circumstances (58,79–82). Therefore, since a good anti-mouse AID antibody is not commercially available for protein detection, we performed quantitative RT-PCR to determine the relative levels of *AID* expression in various mouse tissues. We showed that *AID* is expressed in the spleen (as expected) and ovaries (~60% compared with spleen), but only weakly in heart, and not at all in liver or testes (Figure 8A). These data were in accordance with

Histogram bars represent the mean of three independent cultures, and the standard deviation is shown. WT, wild-type. (B) Western blot showing expression of the HA-tagged proteins from a representative experiment from (A). Tubulin is a loading control. (C) *E. coli*-based Rif<sup>R</sup> mutation assay. See the legend to Figure 3C for assay and label details. (D) Localization of GFP-tagged AID and AID $\Delta C$  in HeLa cells, in the presence and absence of leptomycin B. The scale bars represent 25  $\mu$ M.

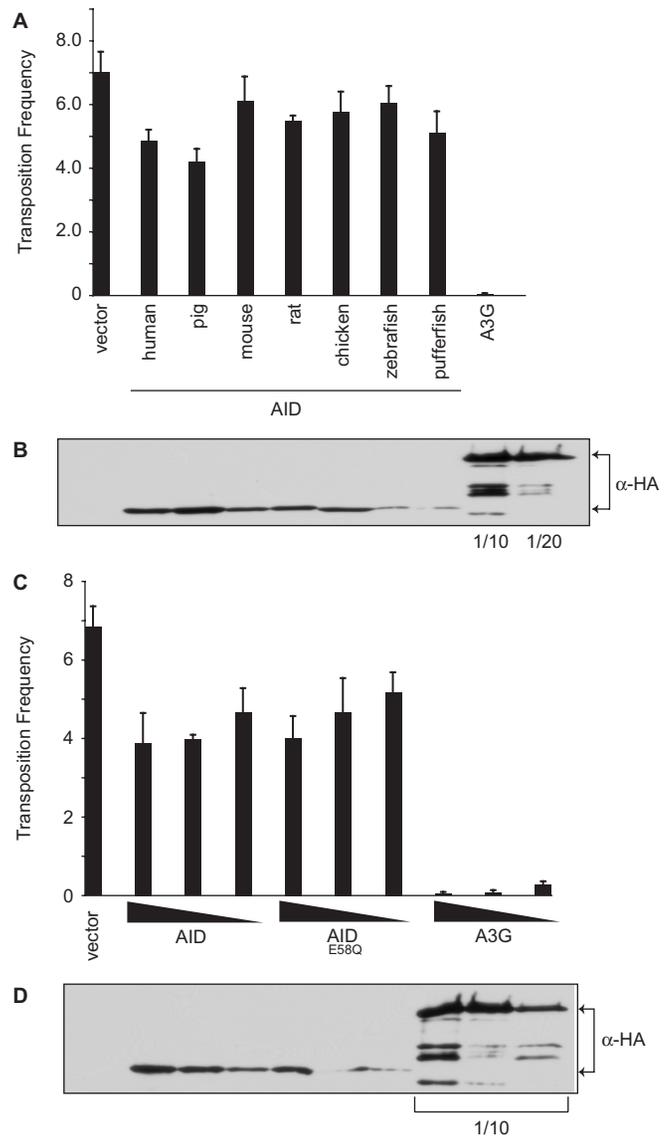


**Figure 5.** Inhibition of L1 transposition by AID phosphorylation site mutants. (A) Percentage of puromycin-resistant cells that were GFP-positive 5 days after transfection with the L1 and indicated AID variants. A3A and A3A<sub>E72A</sub> were included as controls. Histogram bars represent the mean of three independent cultures, and the standard deviation is shown. WT, wild-type. (B) Western blot showing expression of the HA-tagged proteins from a representative experiment from (A). Tubulin is a loading control.

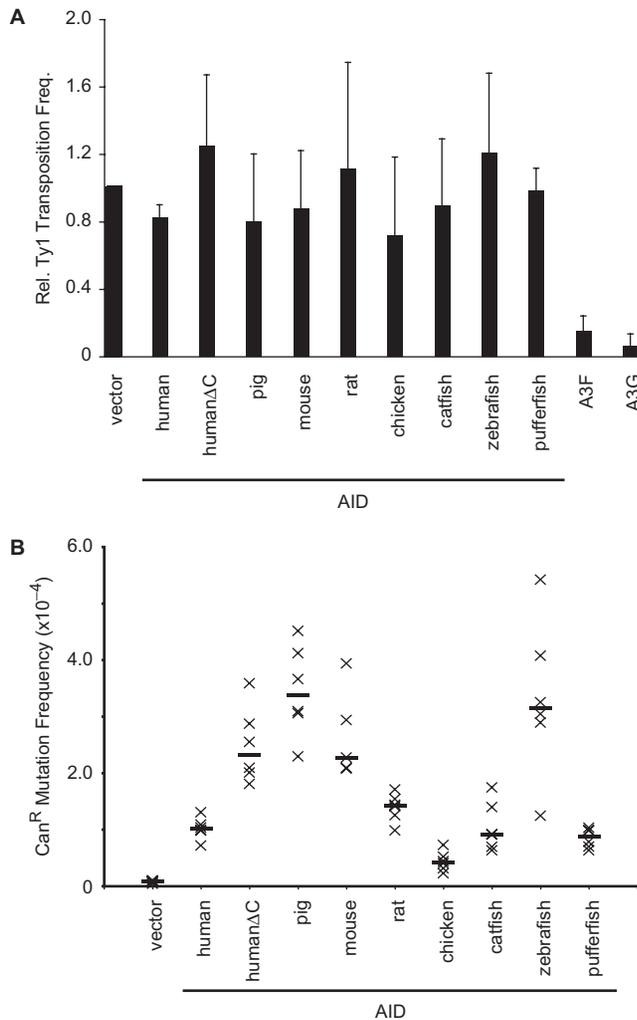
another study that found high levels of *AID* expression in oocytes (82). Representative PCR reactions products were also visualized on an agarose gel to verify the qPCR data (Figure 8B). The AID amplicon was visible in the reaction from spleen and ovary cDNA, but not from heart, liver or testes cDNA. We confirmed specific amplification of AID by cloning and sequencing the qPCR product. *RPL13A* was used as the housekeeping gene control, and it amplified similarly in all samples. We concluded that AID is indeed expressed in a tissue relevant to L1 and MusD replication *in vivo*.

## DISCUSSION

In this study, we showed that AID from multiple species could inhibit the replication of the retrotransposon L1. We were surprised to find that many highly conserved residues of human AID were dispensable for L1 restriction, including the catalytic glutamate E58, the zinc-coordinating cysteines C87 and C90, the active site tryptophan W84 and the single-strand DNA binding arginines R24 and R112. These data combined to demonstrate a DNA deamination-independent mechanism. Similar DNA editing-independent L1 restriction mechanisms have also been documented for human A3B and A3G, but the replicative stage of the block has yet to be determined (22,23,25,31,33,34). In contrast, restriction of L1 by human A3A clearly has a different set of genetic requirements dependent upon the analogous, conserved catalytic glutamate, zinc-coordinating cysteines and active site tryptophan [(25,32) and this study]. Thus, at least two distinct



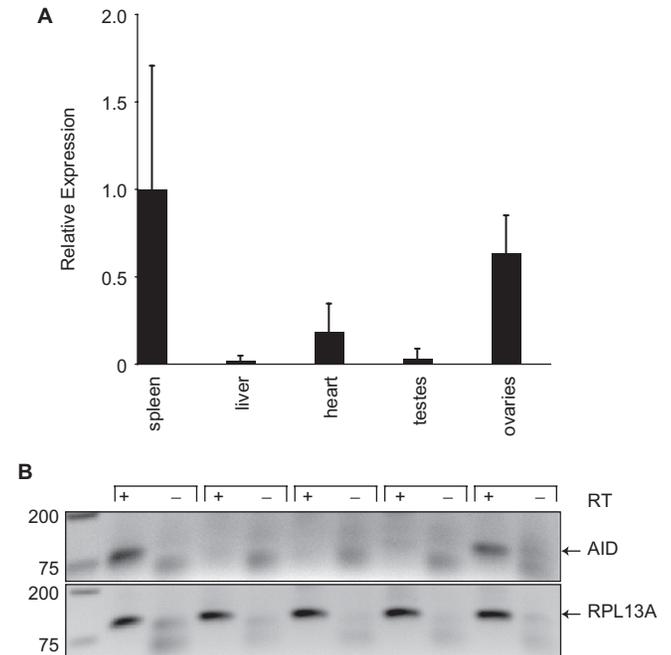
**Figure 6.** Effect of AID on MusD retrotransposition. (A) The transposition frequency of MusD in HeLa cells co-transfected with the MusD plasmid and the indicated AID-HA or A3G-HA expression vectors, and a GFP marker plasmid. Transposition frequency was calculated as the number of neomycin-resistant colonies/transfection efficiency (percent GFP<sup>+</sup>). Histogram bars represent the mean of three independent cultures, and the standard deviation is shown. *P*-values were calculated for each construct using the Student's *t*-test and are as follows: human, *P* = 0.0055; pig, *P* = 0.0020; mouse, *P* = 0.40; rat, *P* = 0.014; chicken, *P* = 0.14; zebrafish, *P* = 0.25; pufferfish, *P* = 0.032; A3G, *P* = 0.000013. (B) Western blot showing expression of the HA-tagged proteins from a representative experiment from (A). The A3G lysate was diluted 1/10 or 1/20, as indicated. (C) The transposition frequency of MusD in HeLa cells co-transfected with the MusD plasmid and varying amounts the AID, AID<sub>E58Q</sub> or A3G expression vectors, and a GFP marker plasmid. The amount of AID/A3G plasmid transfected is indicated below the graph. Transposition frequency is calculated as the number of neomycin-resistant colonies/transfection efficiency (percent GFP<sup>+</sup>). Bars represent the mean of three independent cultures. Error bars indicate the standard deviation from the mean. (D) Western blot showing expression of the HA-tagged proteins from a representative experiment from (C). The A3G lysates were diluted 1/10, as indicated.



**Figure 7.** Effect of AID on Ty1 retrotransposition. **(A)** The transposition frequency of Ty1 in yeast cells expressing the indicated untagged AID or A3 (A3F, A3G) proteins or empty vector. The transposition frequency was calculated as the number of HIS<sup>+</sup> colonies per viable cell. The vector control was normalized to 1 for each experiment. Histogram bars represent the mean of the medians from four independent experiments (actual mean for vector =  $2.8 \times 10^{-6}$ ) and error bars represent the standard deviation. **(B)** Yeast-based CAN<sup>R</sup> assay for DNA mutation. Each 'x' represents the mutation frequency of an independent culture, calculated as the number of CAN<sup>R</sup> colonies per viable cell. Six independent cultures were assayed for each AID variant, and the median mutation frequencies are indicated by the horizontal bars. Vector represents the background level of mutation in *UNG*-deficient yeast.

mechanisms may serve to limit the transposition of L1 and similar non-LTR type retroelements in vertebrates. Additional studies will be required to pinpoint the step (or steps) at which AID/A3 proteins interfere with L1 replication.

Although the restriction of many retrotransposon and retrovirus substrates by multiple A3 proteins is clearly associated with G-to-A hypermutations in the coding DNA strand, a growing number of studies have indicated deamination-independent mechanisms (14,15). Several prior studies showed that L1 restriction was not associated



**Figure 8.** Tissue distribution of AID expression. **(A)** A quantitative analysis of *AID* mRNA expression in mouse spleen, heart, liver, testes and ovaries in relation to *RPL13A* (housekeeping gene control) mRNA in the same samples. Spleen was normalized to 1.0. Tissues from two or three animals were procured and each sample was analyzed in triplicate. The standard deviations are shown. **(B)** Agarose gel images of the PCR products from the qPCR. DNA marker sizes are indicated on the left in base pairs. RT, reverse transcriptase.

with A3 protein subcellular localization, as A3F (cytoplasmic) and A3B (nuclear) each inhibited L1 replication with similar efficiencies (22,23,25,32). Our studies with AID (predominantly cytoplasmic) and AIDΔC (mostly nuclear) also indicated that sub-cellular compartmentalization is not a rate-limiting step in L1 restriction. Thus, we favor an L1 restriction model in which cytoplasmic AID (A3B, A3F or A3G) engages assembling L1 replication complexes co- or post-translationally (Supplementary Figure S7). This cytoplasmic restriction model is supported indirectly by data showing that at least one family member, A3G, forms high-molecular mass cytoplasmic complexes, and that this complex is required for restriction of the L1-dependent retroelement Alu (24,83–85). AID also appears capable of forming large cytoplasmic complexes (1,2,86), and it is possible that at least some of the complex components are shared with A3G. A cytoplasmic restriction model is also appealing because it is the first to suggest a function for AID in the subcellular compartment in which it is predominantly located.

It is notable that G-to-A hypermutations have not been associated with L1 restriction by AID, A3B, A3F, A3G or even A3A, which requires key catalytic residues [(22,23,25,31–34) and this study]. As mentioned above, a part of this apparent failure (at least for A3A) may be due to rapid degradation of edited L1's by cellular DNA repair enzymes. However, a reasonable alternative explanation is the possibility that editing may be co-factor or post-translational modification dependent. For instance, like

the RNA editing family member APOBEC1, C-to-U deamination of its physiological substrate APOB mRNA is thought to require a co-factor called ACF. It is therefore plausible that AID (and also APOBEC3s) employ tissue-specific co-factors and/or post-translational modifications for maximal editing and/or restriction efficiency. AID may very well use distinct co-factors or modifications in germinal center B cells to edit antibody gene DNA than those that it may use in other tissues to inhibit L1 replication (such as the ovary). For instance, cofactors have recently been implicated for AID in antibody diversification and for A3G-dependent restriction of HIV-1 (87,88).

We also confirmed a previous report that AID mRNA is expressed in ovarian tissue (82), at least in mice, placing AID in a compartment where replication of L1 may have the greatest impact *in vivo* (75–77). Although AID expression has also been detected in human testes (81), we did not find it in mouse testes tissue; species-specific differences, sterile housing conditions and/or other factors may be responsible. Nevertheless, expression data [(80–82,89) and this study], L1 restriction data (this study) and phylogenetic evidence that an AID-like ancestor duplicated and diverged to give rise to the present day A3s (41–44) combine to indicate that AID may indeed also function in providing innate immunity to retroelements [as originally proposed in (45); see also (90)]. Future studies will be directed toward testing this hypothesis *in vivo*, in mice or fish or an organism where both precise genetics and quantitative transposition assays are possible.

Endogenous retrotransposons make up a large portion of the mammalian genome, with L1 elements alone constituting around 20% of the human and mouse genomes (91,92). With approximately 100 active copies in humans and around 3000 active copies in mice (93–95), it is likely that cellular factors are required to maintain a balance between new insertions contributing to beneficial genetic diversity and causing deleterious gene disruptions. The APOBEC3 proteins are almost certainly one class of innate restriction factors. It is probable that A3 genes evolved from their more ancient family member, AID, and fine-tuned the anti-retroelement activity during this process. In contrast to many of the APOBEC3s, which appear to have undergone a strong diversifying selection (41,43,96,97), AID appears to be under purifying selection (41), which probably reflects the essential nature of AID's function in antibody diversification. The A3s may have therefore emerged to enable mammals to 'keep up' with the retroviruses and retrotransposons due to AID's functional constraints.

## SUPPLEMENTARY DATA

Supplementary Data is available at NAR Online.

## ACKNOWLEDGEMENTS

We thank B. Cullen, S. Ekker, L. Hammarström, T. Hiedmann, H. Kazazian Jr., N. Landau, D. Largaespada, B. Magor, D. Nissley, M. Nussenzweig, Q. Pan-Hammarström and H. Towle for reagents, the

University of Minnesota Cancer Center Flow Cytometry Core Facility for instrumentation and technical assistance, M. Titus for microscopy facilities, M. Sanders and H. Towle for the use of their homogenizer, M. Stenglein, R. LaRue, K. Silverstein and S. Offer for technical advice and constructs, and W. Brown, J. DiNoia, L. Lackey, H. Malik, H. Matsuo and M. Stenglein for their helpful comments.

## FUNDING

National Institutes of Health (AI064046 and GM090437); University of Minnesota Leukemia Research Fund (to R.S.H.); University of Minnesota Graduate School Doctoral Dissertation Fellowship (to D.A.M.). Funding for open access charge: National Institute of General Medical Sciences, National Institutes of Health, Bethesda, Maryland (GM090437).

*Conflict of interest statement.* None declared.

## REFERENCES

- Dickerson,S.K., Market,E., Besmer,E. and Papavasiliou,F.N. (2003) AID mediates hypermutation by deaminating single stranded DNA. *J. Exp. Med.*, **197**, 1291–1296.
- Chaudhuri,J., Tian,M., Khuong,C., Chua,K., Pinaud,E. and Alt,F.W. (2003) Transcription-targeted DNA deamination by the AID antibody diversification enzyme. *Nature*, **422**, 726–730.
- Bransteitter,R., Pham,P., Scharff,M.D. and Goodman,M.F. (2003) Activation-induced cytidine deaminase deaminates deoxycytidine on single-stranded DNA but requires the action of RNase. *Proc. Natl Acad. Sci. USA*, **100**, 4102–4107.
- Petersen-Mahrt,S.K., Harris,R.S. and Neuberger,M.S. (2002) AID mutates E. coli suggesting a DNA deamination mechanism for antibody diversification. *Nature*, **418**, 99–103.
- Harris,R.S., Sale,J.E., Petersen-Mahrt,S.K. and Neuberger,M.S. (2002) AID is essential for immunoglobulin V gene conversion in a cultured B cell line. *Curr. Biol.*, **12**, 435–438.
- Arakawa,H., Hauschild,J. and Buerstedde,J.M. (2002) Requirement of the activation-induced deaminase (AID) gene for immunoglobulin gene conversion. *Science*, **295**, 1301–1306.
- Revy,P., Muto,T., Levy,Y., Geissmann,F., Plebani,A., Sanal,O., Catalan,N., Forveille,M., Dufourcq-Labelouse,R., Gennery,A. *et al.* (2000) Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2). *Cell*, **102**, 565–575.
- Muramatsu,M., Kinoshita,K., Fagarasan,S., Yamada,S., Shinkai,Y. and Honjo,T. (2000) Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell*, **102**, 553–563.
- Minegishi,Y., Lavoie,A., Cunningham-Rundles,C., Bedard,P.M., Hebert,J., Cote,L., Dan,K., Sedlak,D., Buckley,R.H., Fischer,A. *et al.* (2000) Mutations in activation-induced cytidine deaminase in patients with hyper IgM syndrome. *Clin. Immunol.*, **97**, 203–210.
- Cannon,J.P., Haire,R.N., Rast,J.P. and Litman,G.W. (2004) The phylogenetic origins of the antigen-binding receptors and somatic diversification mechanisms. *Immunol. Rev.*, **200**, 12–22.
- Mussmann,R., Wilson,M., Marcuz,A., Courtet,M. and Du Pasquier,L. (1996) Membrane exon sequences of the three Xenopus Ig classes explain the evolutionary origin of mammalian isotypes. *Eur. J. Immunol.*, **26**, 409–414.
- Greenberg,A.S., Avila,D., Hughes,M., Hughes,A., McKinney,E.C. and Flajnik,M.F. (1995) A new antigen receptor gene family that undergoes rearrangement and extensive somatic diversification in sharks. *Nature*, **374**, 168–173.
- Hinds-Frey,K.R., Nishikata,H., Litman,R.T. and Litman,G.W. (1993) Somatic variation precedes extensive diversification of germline sequences and combinatorial joining in the evolution

- of immunoglobulin heavy chain diversity. *J. Exp. Med.*, **178**, 815–824.
14. Chiu, Y.L. and Greene, W.C. (2008) The APOBEC3 cytidine deaminases: an innate defensive network opposing exogenous retroviruses and endogenous retroelements. *Annu. Rev. Immunol.*, **26**, 317–353.
  15. Malim, M.H. and Emerman, M. (2008) HIV-1 accessory proteins—ensuring viral survival in a hostile environment. *Cell Host Microbe*, **3**, 388–398.
  16. Schumacher, A.J., Haché, G., Macduff, D.A., Brown, W.L. and Harris, R.S. (2008) The DNA deaminase activity of human APOBEC3G is required for Ty1, MusD, and human immunodeficiency virus type 1 restriction. *J. Virol.*, **82**, 2652–2660.
  17. Miyagi, E., Opi, S., Takeuchi, H., Khan, M., Goila-Gaur, R., Kao, S. and Strebel, K. (2007) Enzymatically active APOBEC3G is required for efficient inhibition of human immunodeficiency virus type 1. *J. Virol.*, **81**, 13346–13353.
  18. Zhang, H., Yang, B., Pomerantz, R.J., Zhang, C., Arunachalam, S.C. and Gao, L. (2003) The cytidine deaminase CEM15 induces hypermutation in newly synthesized HIV-1 DNA. *Nature*, **424**, 94–98.
  19. Harris, R.S., Bishop, K.N., Sheehy, A.M., Craig, H.M., Petersen-Mahrt, S.K., Watt, I.N., Neuberger, M.S. and Malim, M.H. (2003) DNA deamination mediates innate immunity to retroviral infection. *Cell*, **113**, 803–809.
  20. Sheehy, A.M., Gaddis, N.C., Choi, J.D. and Malim, M.H. (2002) Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. *Nature*, **418**, 646–650.
  21. Holmes, R.K., Koning, F.A., Bishop, K.N. and Malim, M.H. (2007) APOBEC3F can inhibit the accumulation of HIV-1 reverse transcription products in the absence of hypermutation. Comparisons with APOBEC3G. *J. Biol. Chem.*, **282**, 2587–2595.
  22. Stenglein, M.D. and Harris, R.S. (2006) APOBEC3B and APOBEC3F inhibit L1 retrotransposition by a DNA deamination-independent mechanism. *J. Biol. Chem.*, **281**, 16837–16841.
  23. Muckenfuss, H., Hamdorf, H.L., Held, U., Perkovic, M., Lower, J., Cichutek, K., Flory, E., Schumann, G.G. and Munk, C. (2006) APOBEC3 proteins inhibit human LINE-1 retrotransposition. *J. Biol. Chem.*, **281**, 22161–22172.
  24. Chiu, Y.L., Witkowska, H.E., Hall, S.C., Santiago, M., Soros, V.B., Esnault, C., Heidmann, T. and Greene, W.C. (2006) High-molecular-mass APOBEC3G complexes restrict Alu retrotransposition. *Proc. Natl Acad. Sci. USA*, **103**, 15588–15593.
  25. Bogerd, H.P., Wiegand, H.L., Hulme, A.E., Garcia-Perez, J.L., O’Shea, K.S., Moran, J.V. and Cullen, B.R. (2006) Cellular inhibitors of long interspersed element 1 and Alu retrotransposition. *Proc. Natl Acad. Sci. USA*, **103**, 8780–8785.
  26. Schumacher, A.J., Nissley, D.V. and Harris, R.S. (2005) APOBEC3G hypermutates genomic DNA and inhibits Ty1 retrotransposition in yeast. *Proc. Natl Acad. Sci. USA*, **102**, 9854–9859.
  27. Esnault, C., Heidmann, O., Delebecque, F., Dewannieux, M., Ribet, D., Hance, A.J., Heidmann, T. and Schwartz, O. (2005) APOBEC3G cytidine deaminase inhibits retrotransposition of endogenous retroviruses. *Nature*, **433**, 430–433.
  28. Dutko, J.A., Schafer, A., Kenny, A.E., Cullen, B.R. and Curcio, M.J. (2005) Inhibition of a yeast LTR retrotransposon by human APOBEC3 cytidine deaminases. *Curr. Biol.*, **15**, 661–666.
  29. OhAinle, M., Kerns, J.A., Li, M.M., Malik, H.S. and Emerman, M. (2008) Antiretroelement activity of APOBEC3H was lost twice in recent human evolution. *Cell Host Microbe*, **4**, 249–259.
  30. Niewiadomska, A.M., Tian, C., Tan, L., Wang, T., Sarkis, P.T. and Yu, X.F. (2007) Differential inhibition of long interspersed element 1 by APOBEC3 does not correlate with high-molecular-mass-complex formation or P-body association. *J. Virol.*, **81**, 9577–9583.
  31. Kinomoto, M., Kanno, T., Shimura, M., Ishizaka, Y., Kojima, A., Kurata, T., Sata, T. and Tokunaga, K. (2007) All APOBEC3 family proteins differentially inhibit LINE-1 retrotransposition. *Nucleic Acids Res.*, **35**, 2955–2964.
  32. Chen, H., Lilley, C.E., Yu, Q., Lee, D.V., Chou, J., Narvaiza, I., Landau, N.R. and Weitzman, M.D. (2006) APOBEC3A is a potent inhibitor of adeno-associated virus and retrotransposons. *Curr. Biol.*, **16**, 480–485.
  33. Hulme, A.E., Bogerd, H.P., Cullen, B.R. and Moran, J.V. (2007) Selective inhibition of Alu retrotransposition by APOBEC3G. *Gene*, **390**, 199–205.
  34. Turelli, P., Vianin, S. and Trono, D. (2004) The innate antiretroviral factor APOBEC3G does not affect human LINE-1 retrotransposition in a cell culture assay. *J. Biol. Chem.*, **279**, 43371–43373.
  35. Esnault, C., Millet, J., Schwartz, O. and Heidmann, T. (2006) Dual inhibitory effects of APOBEC family proteins on retrotransposition of mammalian endogenous retroviruses. *Nucleic Acids Res.*, **34**, 1522–1531.
  36. Esnault, C., Priet, S., Ribet, D., Heidmann, O. and Heidmann, T. (2008) Restriction by APOBEC3 proteins of endogenous retroviruses with an extracellular life cycle: ex vivo effects and in vivo “traces” on the murine IAP and human HERV-K elements. *Retrovirology*, **5**, 75.
  37. Jern, P., Stoye, J.P. and Coffin, J.M. (2007) Role of APOBEC3 in genetic diversity among endogenous murine leukemia viruses. *PLoS Genet.*, **3**, 2014–2022.
  38. Lee, Y.N., Malim, M.H. and Bieniasz, P.D. (2008) Hypermutation of an ancient human retrovirus by APOBEC3G. *J. Virol.*, **82**, 8762–8770.
  39. Cost, G.J., Feng, Q., Jacquier, A. and Boeke, J.D. (2002) Human L1 element target-primed reverse transcription in vitro. *EMBO J.*, **21**, 5899–5910.
  40. Luan, D.D., Korman, M.H., Jakubczak, J.L. and Eickbush, T.H. (1993) Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition. *Cell*, **72**, 595–605.
  41. LaRue, R.S., Jónsson, S.R., Silverstein, K.A.T., Lajoie, M., Bertrand, D., El-Mabrouk, N., Hötzel, I., Andrésdóttir, V., Timothy, P.L., Smith, T.P.L. et al. (2008) The artiodactyl APOBEC3 innate immune repertoire shows evidence for a multi-functional domain organization that existed in the ancestor of placental mammals. *BMC Mol. Biol.*, **9**, 104.
  42. Conticello, S.G., Langlois, M.A., Yang, Z. and Neuberger, M.S. (2007) DNA deamination in immunity: AID in the context of its APOBEC relatives. *Adv. Immunol.*, **94**, 37–73.
  43. Conticello, S.G., Thomas, C.J., Petersen-Mahrt, S.K. and Neuberger, M.S. (2005) Evolution of the AID/APOBEC family of polynucleotide (deoxy)cytidine deaminases. *Mol. Biol. Evol.*, **22**, 367–377.
  44. Harris, R.S. and Liddament, M.T. (2004) Retroviral restriction by APOBEC proteins. *Nat. Rev. Immunol.*, **4**, 868–877.
  45. Macduff, D.A. and Harris, R.S. (2006) Directed DNA deamination by AID/APOBEC3 in immunity. *Curr Biol*, **16**, R186–R189.
  46. Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R. et al. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics*, **23**, 2947–2948.
  47. Suyama, M., Torrents, D. and Bork, P. (2006) PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Res.*, **34**, W609–W612.
  48. Clamp, M., Cuff, J., Searle, S.M. and Barton, G.J. (2004) The Jalview Java alignment editor. *Bioinformatics*, **20**, 426–427.
  49. Kosakovsky Pond, S.L., Posada, D., Gravenor, M.B., Woelk, C.H. and Frost, S.D. (2006) GARD: a genetic algorithm for recombination detection. *Bioinformatics*, **22**, 3096–3098.
  50. Retief, J.D. (2000) Phylogenetic analysis using PHYLIP. *Methods Mol. Biol.*, **132**, 243–258.
  51. Uenishi, H., Eguchi, T., Suzuki, K., Sawazaki, T., Toki, D., Shinkai, H., Okumura, N., Hamasima, N. and Awata, T. (2004) PEDE (Pig EST Data Explorer): construction of a database for ESTs derived from porcine full-length cDNA libraries. *Nucleic Acids Res.*, **32**, D484–D488.
  52. Buerstedde, J.M. and Takeda, S. (1991) Increased ratio of targeted to random integration after transfection of chicken B cell lines. *Cell*, **67**, 179–188.
  53. Yu, Q., Chen, D., Konig, R., Mariani, R., Unutmaz, D. and Landau, N.R. (2004) APOBEC3B and APOBEC3C are potent inhibitors of simian immunodeficiency virus replication. *J. Biol. Chem.*, **279**, 53379–53386.

54. MacDuff, D.A., Neuberger, M.S. and Harris, R.S. (2006) MDM2 can interact with the C-terminus of AID but it is inessential for antibody diversification in DT40 B cells. *Mol. Immunol.*, **43**, 1099–1108.
55. Harris, R.S., Petersen-Mahrt, S.K. and Neuberger, M.S. (2002) RNA editing enzyme APOBEC1 and some of its homologs can act as DNA mutators. *Mol. Cell.*, **10**, 1247–1253.
56. Ichikawa, H.T., Sowden, M.P., Torelli, A.T., Bachl, J., Huang, P., Dance, G.S., Marr, S.H., Robert, J., Wedekind, J.E., Smith, H.C. *et al.* (2006) Structural phylogenetic analysis of activation-induced deaminase function. *J. Immunol.*, **177**, 355–361.
57. Barreto, V.M., Pan-Hammarström, Q., Zhao, Y., Hammarström, L., Misulovin, Z. and Nussenzweig, M.C. (2005) AID from bony fish catalyzes class switch recombination. *J. Exp. Med.*, **202**, 733–738.
58. Saunders, H.L. and Magor, B.G. (2004) Cloning and expression of the AID gene in the channel catfish. *Dev. Comp. Immunol.*, **28**, 657–663.
59. Ostertag, E.M., Prak, E.T., DeBerardinis, R.J., Moran, J.V. and Kazazian, H.H. Jr (2000) Determination of L1 retrotransposition kinetics in cultured cells. *Nucleic Acids Res.*, **28**, 1418–1423.
60. Chen, K.M., Harjes, E., Gross, P.J., Fahmy, A., Lu, Y., Shindo, K., Harris, R.S. and Matsuo, H. (2008) Structure of the DNA deaminase domain of the HIV-1 restriction factor APOBEC3G. *Nature*, **452**, 116–119.
61. Rada, C., Jarvis, J.M. and Milstein, C. (2002) AID-GFP chimeric protein increases hypermutation of Ig genes with no evidence of nuclear localization. *Proc. Natl Acad. Sci. USA*, **99**, 7003–7008.
62. McBride, K.M., Barreto, V., Ramiro, A.R., Stavropoulos, P. and Nussenzweig, M.C. (2004) Somatic hypermutation is limited by CRM1-dependent nuclear export of activation-induced deaminase. *J. Exp. Med.*, **199**, 1235–1244.
63. Ito, S., Nagaoka, H., Shinkura, R., Begum, N., Muramatsu, M., Nakata, M. and Honjo, T. (2004) Activation-induced cytidine deaminase shuttles between nucleus and cytoplasm like apolipoprotein B mRNA editing catalytic polypeptide 1. *Proc. Natl Acad. Sci. USA*, **101**, 1975–1980.
64. Pasqualucci, L., Kitaura, Y., Gu, H. and Dalla-Favera, R. (2006) PKA-mediated phosphorylation regulates the function of activation-induced deaminase (AID) in B cells. *Proc. Natl Acad. Sci. USA*, **103**, 395–400.
65. Basu, U., Chaudhuri, J., Alpert, C., Dutt, S., Ranganath, S., Li, G., Schrum, J.P., Manis, J.P. and Alt, F.W. (2005) The AID antibody diversification enzyme is regulated by protein kinase A phosphorylation. *Nature*, **438**, 508–511.
66. McBride, K.M., Gazumyan, A., Woo, E.M., Schwickert, T.A., Chait, B.T. and Nussenzweig, M.C. (2008) Regulation of class switch recombination and somatic mutation by AID phosphorylation. *J. Exp. Med.*, **205**, 2585–2594.
67. Chatterji, M., Unniraman, S., McBride, K.M. and Schatz, D.G. (2007) Role of activation-induced deaminase protein kinase A phosphorylation sites in Ig gene conversion and somatic hypermutation. *J. Immunol.*, **179**, 5274–5280.
68. McBride, K.M., Gazumyan, A., Woo, E.M., Barreto, V.M., Robbiani, D.F., Chait, B.T. and Nussenzweig, M.C. (2006) Regulation of hypermutation by activation-induced cytidine deaminase phosphorylation. *Proc. Natl Acad. Sci. USA*, **103**, 8798–8803.
69. Bogerd, H.P., Tallmadge, R.L., Oaks, J.L., Carpenter, S. and Cullen, B.R. (2008) Equine infectious anemia virus resists the anti-retroviral activity of equine APOBEC3 proteins through a packaging-independent mechanism. *J. Virol.*, **82**, 11889–11901.
70. Bogerd, H.P., Wiegand, H.L., Doehle, B.P., Lueders, K.K. and Cullen, B.R. (2006) APOBEC3A and APOBEC3B are potent inhibitors of LTR-retrotransposon function in human cells. *Nucleic Acids Res.*, **34**, 89–95.
71. Armitage, A.E., Katzourakis, A., de Oliveira, T., Welch, J.J., Belshaw, R., Bishop, K.N., Kramer, B., McMichael, A.J., Rambaut, A. and Iversen, A.K. (2008) Conserved footprints of APOBEC3G on hypermutated human immunodeficiency virus type 1 and human endogenous retrovirus HERV-K(HML2) sequences. *J. Virol.*, **82**, 8743–8761.
72. Jónsson, S.R., LaRue, R.S., Stenglein, M.D., Fahrenkrug, S.C., Andrésdóttir, V. and Harris, R.S. (2007) The restriction of zoonotic PERV transmission by human APOBEC3G. *PLoS ONE*, **2**, e893.
73. Ribet, D., Dewannieux, M. and Heidmann, T. (2004) An active murine transposon family pair: retrotransposition of “master” MusD copies and ETn trans-mobilization. *Genome Res.*, **14**, 2261–2267.
74. Poltoratsky, V.P., Wilson, S.H., Kunkel, T.A. and Pavlov, Y.I. (2004) Recombinogenic phenotype of human activation-induced cytosine deaminase. *J. Immunol.*, **172**, 4308–4313.
75. Prak, E.T., Dodson, A.W., Farkash, E.A. and Kazazian, H.H. Jr (2003) Tracking an embryonic L1 retrotransposition event. *Proc. Natl Acad. Sci. USA*, **100**, 1832–1837.
76. Ostertag, E.M., DeBerardinis, R.J., Goodier, J.L., Zhang, Y., Yang, N., Gerton, G.L. and Kazazian, H.H. Jr (2002) A mouse model of human L1 retrotransposition. *Nat. Genet.*, **32**, 655–660.
77. Brouha, B., Meischl, C., Ostertag, E., de Boer, M., Zhang, Y., Neijens, H., Roos, D. and Kazazian, H.H. Jr (2002) Evidence consistent with human L1 retrotransposition in maternal meiosis I. *Am. J. Hum. Genet.*, **71**, 327–336.
78. Maksakova, I.A., Romanish, M.T., Gagnier, L., Dunn, C.A., van de Lagemaat, L.N. and Mager, D.L. (2006) Retroviral elements and their hosts: insertional mutagenesis in the mouse germ line. *PLoS Genet.*, **2**, e2.
79. Matsumoto, Y., Marusawa, H., Kinoshita, K., Endo, Y., Kou, T., Morisawa, T., Azuma, T., Okazaki, I.M., Honjo, T. and Chiba, T. (2007) Helicobacter pylori infection triggers aberrant expression of activation-induced cytidine deaminase in gastric epithelium. *Nat. Med.*, **13**, 470–476.
80. Endo, Y., Marusawa, H., Kinoshita, K., Morisawa, T., Sakurai, T., Okazaki, I.M., Watashi, K., Shimotohno, K., Honjo, T. and Chiba, T. (2007) Expression of activation-induced cytidine deaminase in human hepatocytes via NF-kappaB signaling. *Oncogene*, **26**, 5587–5595.
81. Schreck, S., Buettner, M., Kremmer, E., Bogdan, M., Herbst, H. and Niedobitek, G. (2006) Activation-induced cytidine deaminase (AID) is expressed in normal spermatogenesis but only infrequently in testicular germ cell tumours. *J. Pathol.*, **210**, 26–31.
82. Morgan, H.D., Dean, W., Coker, H.A., Reik, W. and Petersen-Mahrt, S.K. (2004) Activation-induced cytidine deaminase deaminates 5-methylcytosine in DNA and is expressed in pluripotent tissues: implications for epigenetic reprogramming. *J. Biol. Chem.*, **279**, 52353–52360.
83. Wang, X., Dolan, P.T., Dang, Y. and Zheng, Y.H. (2007) Biochemical differentiation of APOBEC3F and APOBEC3G proteins associated with HIV-1 life cycle. *J. Biol. Chem.*, **282**, 1585–1594.
84. Kreisberg, J.F., Yonemoto, W. and Greene, W.C. (2006) Endogenous factors enhance HIV infection of tissue naive CD4 T cells by stimulating high molecular mass APOBEC3G complex formation. *J. Exp. Med.*, **203**, 865–870.
85. Chiu, Y.L., Soros, V.B., Kreisberg, J.F., Stopak, K., Yonemoto, W. and Greene, W.C. (2005) Cellular APOBEC3G restricts HIV-1 infection in resting CD4+ T cells. *Nature*, **435**, 108–114.
86. Ta, V.T., Nagaoka, H., Catalan, N., Durandy, A., Fischer, A., Imai, K., Nonoyama, S., Tashiro, J., Ikegawa, M., Ito, S. *et al.* (2003) AID mutant analyses indicate requirement for class-switch-specific cofactors. *Nat. Immunol.*, **4**, 843–848.
87. Conticello, S.G., Ganesh, K., Xue, K., Lu, M., Rada, C. and Neuberger, M.S. (2008) Interaction between antibody-diversification enzyme AID and spliceosome-associated factor CTNBL1. *Mol. Cell*, **31**, 474–484.
88. Han, Y., Wang, X., Dang, Y. and Zheng, Y. H. (2008) APOBEC3G and APOBEC3F require an endogenous cofactor to block HIV-1 replication. *PLoS Pathog.*, **4**, e1000095.
89. Kou, T., Marusawa, H., Kinoshita, K., Endo, Y., Okazaki, I.M., Ueda, Y., Kodama, Y., Haga, H., Ikai, I. and Chiba, T. (2007) Expression of activation-induced cytidine deaminase in human hepatocytes during hepatocarcinogenesis. *Int. J. Cancer.*, **120**, 469–476.
90. Gourzi, P., Leonova, T. and Papavasiliou, F.N. (2006) A role for activation-induced cytidine deaminase in the host response against a transforming retrovirus. *Immunity*, **24**, 779–786.

91. Waterston,R.H., Lindblad-Toh,K., Birney,E., Rogers,J., Abril,J.F., Agarwal,P., Agarwala,R., Ainscough,R., Alexandersson,M., An,P. *et al.* (2002) Initial sequencing and comparative analysis of the mouse genome. *Nature*, **420**, 520–562.
92. Lander,E.S., Linton,L.M., Birren,B., Nusbaum,C., Zody,M.C., Baldwin,J., Devon,K., Dewar,K., Doyle,M., FitzHugh,W. *et al.* (2001) Initial sequencing and analysis of the human genome. *Nature*, **409**, 860–921.
93. Brouha,B., Schustak,J., Badge,R.M., Lutz-Prigge,S., Farley,A.H., Moran,J.V. and Kazazian,H.H. Jr (2003) Hot L1s account for the bulk of retrotransposition in the human population. *Proc. Natl Acad. Sci. USA*, **100**, 5280–5285.
94. Goodier,J.L., Ostertag,E.M., Du,K. and Kazazian,H.H. Jr (2001) A novel active L1 retrotransposon subfamily in the mouse. *Genome Res.*, **11**, 1677–1685.
95. DeBerardinis,R.J., Goodier,J.L., Ostertag,E.M. and Kazazian,H.H. Jr (1998) Rapid amplification of a retrotransposon subfamily is evolving the mouse genome. *Nat. Genet.*, **20**, 288–290.
96. Zhang,J. and Webb,D.M. (2004) Rapid evolution of primate antiviral enzyme APOBEC3G. *Hum. Mol. Genet.*, **13**, 1785–1791.
97. Sawyer,S.L., Emerman,M. and Malik,H.S. (2004) Ancient adaptive evolution of the primate antiviral DNA-editing enzyme APOBEC3G. *PLoS Biol.*, **2**, E275.